Part 2: The Phase-oriented Advice and Review Structure (PARS) for Digital Forensic Investigations (Peer reviewed and accepted 21st September 2020 Forensic Science International: Di...

Article *in* Forensic Science International Digital Investigation · September 2020

CITATIONS
0

READS
94

2 authors, including:

Nina Sunde
Norwegian Police University College, Oslo, Norway
9 PUBLICATIONS **19** CITATIONS

SEE PROFILE

Part 2: The Phase-oriented Advice and Review Structure (PARS) for Digital Forensic Investigations.

Nina Sunde[a*] and Dr Graeme Horsman[b]

[a] Norwegian Police University College, Pb. 2109 Vika, 0125 Oslo, Norway

[b] Teesside University, Middlesbrough, Tees Valley, TS1 3BX, UK

*Corresponding author

Email addresses nina.sunde@phs.no (N. Sunde) g.horsman@tees.ac.uk (G. Horsman)

Abstract

This work forms the second part of a two part series providing the necessary scaffolding for the digital forensic discipline to conduct effective peer review in their laboratories and units. The first part articulated the need for a structured approach to peer review in digital forensic investigations (Horsman and Sunde, 2020). Here in part two, the Phase-oriented Advice and Review Structure (PARS) for digital forensic investigations is offered. PARS is the first documented peer review methodology for the digital forensics field, a six staged approach designed to formally support organisations and their staff in their goal of facilitating effective peer review of DF work, from investigative tasks to forensic activities and forensic analysis processes (Pollitt et al., 2018). This article discusses how the PARS methodology can be implemented, and the available options and mechanisms available to ease the interpretation of this model into existing practices. Both the early 'Advisor' and later 'Reviewer' roles in PARS are discussed and their requirements and expectations are defined. Three template documents are provided and explained: The PARS Advisors template, the PARS Advisor Brief template and the PARS Peer Review Hierarchy template, for direct use by organisations seeking to adopt the PARS methodology.

Keywords: Digital Forensics; Peer Review; Digital Evidence; Quality Assurance; Quality Control; Forensic Science

1 Introduction

This work is the second part offered by the authors covering the topic of peer review for the field of digital forensic (DF) and their practitioners. In Part 1 (Horsman and Sunde, 2020), the need for a structured approach to peer review in the field of DF was articulated along with underpinning literature, highlighting current challenges and issues surrounding peer reviewing DF investigations. Here in Part 2, we move to describe '*what a peer review should look like*' for practitioners working in this discipline and propose the **P**hase-oriented **A**dvice and **R**eview **S**tructure (PARS) for Digital Forensic investigations. PARS is a peer review methodology designed for rollout across the DF discipline to support organisations in the task of implementing a system for reviewing practitioner work in order to uphold and maintain acceptable quality standards. PARS is aimed at supporting any type of DF organisation from DF laboratory environments or smaller DF units, private or public. The structure of PARS is described in detail and the requirements of each component mapped, which facilitates a partial or stepwise implementation. Further, the practicalities and organizational challenges of implementing PARS are considered and discussed, and PARS templates are provided to help practitioners carry out the PARS review process in their workplace.

2 What should peer review look like?

Arguably, a '*traditional*' styled peer review process in DF could be considered a singular entity, which takes place in the closing stage of an investigation. In essence, peer review is often thought of as the final (and in some cases, primary) quality control (QC) check undertaken by an organisation. As a result, it is naturally considered as a single entity; a process of checking 'everything' a practitioner has done following the close of their work on a given case. The problem with this approach is arguably threefold:

1. *Efficiency... or Inefficiency*: To understand the impact that a traditional styled peer review has upon efficiency, one has to consider a peer review, which uncovers significant flaws in a given digital forensic investigation process. In such cases, further fundamental work may be required, which could include the completion of additional (or re-running of) processes or the undertaking of supplementary testing. The problem this causes lies with establishing such issues at the close of an investigation, at a time where a practitioners 'costed' time may already have been reached. Traditional peer review approaches provide limited room to maneuver, as the practitioner has already undertaken all of the work they expected to undertake. Therefore when issues are uncovered, wholesale changes may

be required, which could result in '*overspends*' in terms of time and resources allocated to a specific case. When considered in terms of efficiency/inefficiency, if errors are considered to cause resource wastage (practitioner time etc.), then their mitigation at the earliest opportunity should be considered a priority. In doing so, a potential '*cascading'* and/or '*snowball effect*' of errors may be prevented (Dror et al., 2017). Consider a practitioner who recovers Internet history records, reviews and then reports them, only to find that the process they have used has only provided a partial set of data for the practitioner to examine. In this case the practitioner must start the investigation again to ensure comprehensive data recovery takes place via re-ran processes, effectively duplicating work and doubling the resources spent on the case.

2. *Reactive rather than preventative*: Given that traditional peer reviews take place at the end of the DF process, they are by their very nature a reactive process. They are designed to evaluate in their entirety the complete investigative process and everything that has been generated as a result. In comparison, a peer review which occurs earlier in the investigative process, and at defined stages, can rectify any apparent error earlier, and in some instances prevent any error from impacting on further aspects of a case (as noted above). Perhaps the clearest example of this issue would be where there may be an apparent issue with the acquisition of data from a device, yet a practitioner continues to rely on this dataset in belief that it is the best acquisition of data that could be attained. A traditional peer review might reveal an issue with their acquisition (incomplete, errored sectors or verification issue) or an alternative method that could have extracted more data (for example, logical vs file system mobile device extraction methods). In either case, the practitioner may be required to go back and restart the investigation. Detecting this via a peer review at an earlier stage may prevent such issues. Performing peer review earlier, at all stages of the investigatory process, will give an opportunity to learn and improve skills during casework, instead of learning from mistakes and having to react. This approach also provides an opportunity to identify and correct systematic errors that may be overlooked or harder to detect at a later stage.

3. *Too much to review*: A traditional peer review must evaluate all of the investigatory process that has taken place. This raises the question as to whether this is too big of a task to undertake in one sitting, where a divide and conquer approach to the peer review process

allows more manageable sub-reviews to take place, arguably increasing the chance of identifying errors.

The PARS approach presented in this work aligns the peer review process to the typical stages of a DF investigation, widely documented in academic literature (Köhn et al., 2006; Casey, 2011; Agarwal et al., 2011; Jafari and Satti, 2015). Fig. 1 shows the peer review process staged across the DF investigation process. PARS is inspired by the procedure for periodic review of investigations, described in the ACPO Murder Investigation Manual (2006), and several of the force-level policy documents by National Centre for Policing Excellence (2005 etc. - see also Savage and Milne 2011).

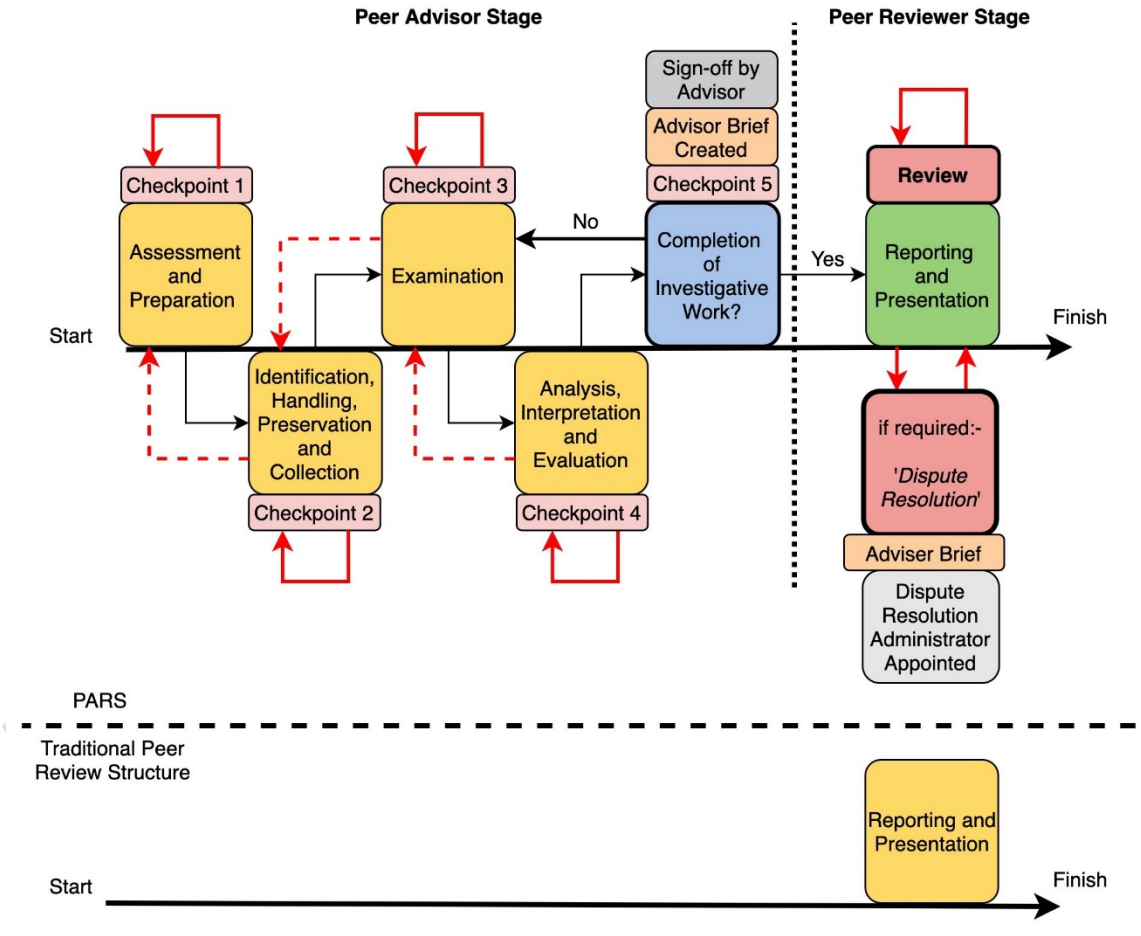


Figure 1: PARS Vs. Traditional peer review models.

PARS makes two fundamental changes to traditional peer review: (1) it is multi-staged with the requirement for 'dispute resolution' procedures, and (2) multi-person divided into two roles (*'Advisor*' and the *'Reviewer'*).

2.1 A multi-staged review

The proposed review structure consists of five 'stages', enforcing an iterative approach to peer review. We propose that each investigation should pass through four '*Advisor Checkpoints*' and a final '*Review*' (discussed below). This approach closely aligns to the main steps a practitioner will address when conducting casework, which are often considered critical milestones in all investigatory processes. In doing so, the review is a constant system of checks and verification to prevent not only a singular error from occurring, but to then stop any such error from impacting further investigative work. The multi-stages approach is designed to lessen the burden of the review by compartmentalising the review into more frequent, but arguably manageable stages.

There is a need to formally acknowledge the need for a 'dispute resolution stage' in the peer review process. There may be the assumption that peer review is a process of review and acceptance, where in reality, narratives may not result in a situation of agreement between the Reviewer(s) and the practitioner. Where disagreement occurs, effective dispute resolution is required in order to allow a review process to conclude, preventing a state of deadlock from occurring. PARS formally acknowledges dispute resolution as a stage of the review process, and a process for dispute resolution is outlined in Section 3.2.2.

2.2 A multi-person review

Whilst counter arguments to such a proposal will lie with resourcing concerns, it is argued that an effective review must be one which is undertaken through multiple agents. Here, a proposal is made to divide the peer review burden between two entities, one who supplements and guides the primary investigator through their casework, followed by a second individual who is independent to the investigative process, who reviews the case in its entirety, often via a review of the written report. Dividing the task between two roles is done for two main reasons. First, advising through Checkpoints and performing peer review of the final report would be a substantial workload for a single person, with a risk of reduced quality due to a too burdensome process. Second, separating the advice and review tasks is justified with the risk of cognitive bias. If the same person should give advice and perform peer review, they would not review with 'fresh

eyes', and would be biased from what they already knew about the DF investigation of the case. Since they already have invested effort to enhance the quality of the result through the peer advice stage, a risk of irrational escalation of commitment (Staw, 1981), which may reduce the ability to sufficiently critical during the peer review stage. In essence, we are proposing that the roles of 'Advisor' and 'Reviewer' (which we introduce in detail below) are two separate individuals.

The 'Advisor' and 'Reviewer' roles: The distinction between the 'Advisor' and 'Reviewer' roles is that of criticality and position in the review process. During the early stages of the PARS review (1-4), those engaging in the peer review process are doing so as a '*critical friend*'. Their role is that of *advice-giver*, taking into account the facts of a given case and the approach of the practitioner, offering recommendations for approaching their tasks and where necessary steering the investigative process. Whilst those advising at each stage will still check critical facts and processes (seen predominantly at stages 1 & 2), the role is to feed-forward, where advice should look to increase the comprehensiveness of the investigation. The Advisor role is proposed to be undertaken by one individual who can guide the practitioner through each of the five Checkpoints (see Fig. 1).

The Advisor role in a DF context is not completely novel, and two roles have been described; an investigative advisor with operational experience, and a forensic advisor with scientific background (Casey et al., 2019, Casey 2019). The advisory roles are designed for harmonization and knowledge management, and should "negotiate the borders between police, DF specialists, criminal intelligence analysts and attorneys to cultivate a criminal justice system that treats digital traces effectively, has visibility across criminal activities, and addresses crime and security more strategically" (Casey, 2019, p.1). Forensic advisors at the Belgian National Institute for Criminalistics and Criminology are generalists in forensic science and act as advising body to the magistrate, and the aim is to improve communication between the magistrate, the police investigators, forensic experts and front-line forensic practitioners (Bitzer et al., 2018). In PARS, the role of the Advisor is different. While ensuring a scientifically sound handling of digital evidence is crucial, the scope of the PARS Advisor tasks may be defined as a combination of investigatory and forensic tasks, since they are meant to oversee a sound handling of the digital evidence through the individual phases of the DF process, and the use of a sound scientific approach to the examination. In contrast to the forensic advisor role as coordinator, the DF Advisor role in PARS is related to supervision and quality assurance (QA).

Advisor Brief: It is argued that this position is based on balanced-efficiency, as having a different Advisor for each stage may become too burdensome for an organisation. However, it is a requirement that the '*Reviewer*' be separate to this process in order to perform an independent, impartial review. This entails that the Reviewer should first conduct the peer review and reach a conclusion (if a peer review on level 4-6 in Peer Review Hierarchy is conduced) based solely on the report and without any knowledge of challenges during the Checkpoints. If the Reviewers conclusion differ from the original conclusion, the Reviewer should then be provided with the Advisor Brief, which is used as a supplementary documentation during dispute resolution (see Section 3.2.2).

3. The PARS model

The PARS model is aimed at all organisations carrying out DF investigatory work, encompassing both private and public sector DF laboratories and DF units. Typically, most DF casework will follow a standardised investigative methodology, but often this methodology omits recognition of the need to peer review. Therefore as a first requirement, organisations must ensure that peer review is formally acknowledged in their existing practice models and given the time and resources to effectively undertake this practice.

Extending the traditional DF investigative model: Over the last 15 years there have been multiple investigative models proposed for DF (see scoping reviews from Kohn et al., (2013) and Du et al.'s (2017) highlighting many of those in existence), and many maintain at their core, the same main aspects of an investigation. Whilst not all (see, for example the acknowledgement of a review stage by Agarwal et al., (2011)), many do not explicitly recognise the requirement for a stage involving robust peer review as part of the DF investigatory process, or omit to discuss the implementation or requirements of it. As a result, this work opts to consolidate existing framework discussions and offers a clear requirement for peer review to be part of a 7 stage DF investigative process, shown in Fig. 2.

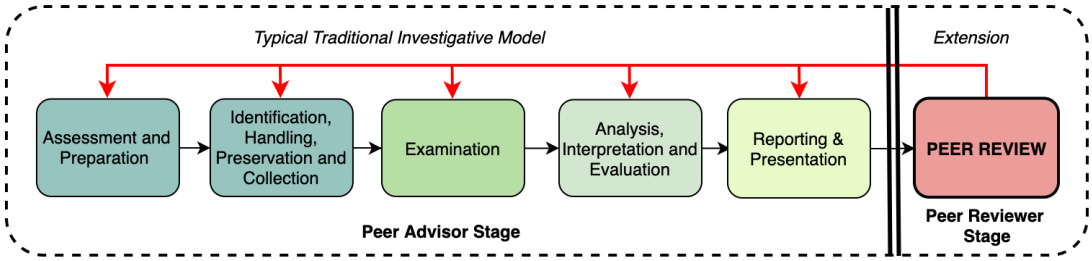


Figure 2: The proposed investigative model incorporating peer review.

Whilst some laboratories may be subject to oversight and governance from certain accreditation standards (for example, in England and Wales, laboratories are expected to now hold the ISO17025 accreditation (Horsman, 2019)), there are also those which operate unregulated. This may be due to a lack of resources at the laboratories disposal in order to attempt to achieve and maintain any accrediting standards, or that cost has effectively ruled it out. The selection of standard is primarily based on the purpose of implementing it, and the deliverable promised from use, which implies that more than one standard is necessary for process consistency and quality assurance (Cusack, 2019). As stated by Cusack (2019, p. 4) “For example, the ISO/IEC 17025 is relevant for the certification of general laboratories, it is considered insufficient to certify digital forensic practice, practitioners, or information security.” Yet Tully et al. (2020) demonstrated that accreditation in itself would not safeguard quality. The accredited forensic laboratories have defined procedures for all the processes, and the practitioners are mandated to follow these. The level of quality may thus not be measured on the defined procedures, but relies on the compliance with these.

Yet regardless of a laboratories position, PARS is a publicly accessible resource to support quality management and can be adopted either in line with, or external to, any accreditation regime if a DF laboratory or unit does not comply with one. It is important to note that PARS can offer support to all DF laboratories and units, regardless of their operational practices as it is designed to bolster peer review.

3.1 The Peer Advisor Stage

If we consider PARS as a compartmentalised process, each element depicted in Fig. 1 above must be examined. A checklist for each component is provided in Appendix 1, 2 and 3.

3.1.1 Checkpoint 1: Assessment and preparation

Checkpoint 1 considers all content prior to commencing an actual investigative process. Here, the Advisor should cover five main procedural branches, supported by the checklist provided in Appendix 1.

1. *Case paperwork and data set verification*: Whilst this may in practice be a simple task, it is important to ensure this is done at the earliest stage possible. Verification of the case paperwork should be undertaken to determine the correct point of contact for an associated client(s), whether there is an agreed contract of work, and whether the remit of work is correct.

   If exhibits/data sets are received, they should be assessed/controlled:

   a. All exhibits are present and verifiable with associated paperwork, documenting a full and sustained chain of custody.
   b. All exhibits are relevant and suitable for the aim of the task.
   c. All exhibits are complete, accurate and authentic. Any deviations should be documented accurately.

2. *Client requirements and expectations*: An evaluation of client expectations and requirements is needed, doubling as a 'feasibility assessment' of work to be done. It is important that at the start of any case, the client is fully and accurately aware of the expected product of the work to be done. This is closely linked with the factors of cost and time frame, as both can impact upon what can be feasibly achieved within the assumed agreed time of an investigation. Establishing the relationship and agreement between client and investigator, and ensuring that expectations are accurately defined, is important to preventing the need for any dispute resolution in the latter stages of the investigative process if expectations have not been met. It is also important to clarify that the practitioner and their lab possess the agreed capability for meeting the expectations of the client. For example, ensuring that a client is aware of any technical limitations and refraining from promising to deliver on a task, which may not be achievable.

3. *The task/mandate and case information:* The outcome of the DF process relies very much on how the initial task or mandate is described, and whether the practitioner receives the

necessary information. Starting an analysis with a general and broad “find evidence” approach does not necessarily lead to either quality or efficiency. One could argue that the party describing the task or mandate is responsible for it being accurate and clear. However, this process should be seen as a dual responsibility. The party requiring the analysis should be responsible for conveying the task-relevant case information (Dror and Cole, 2010; National Academy of Sciences, 2009; President's Council of Advisors on Science and Technology, 2016; UK Forensic Science Regulator, 2015), and the DF practitioner should ensure that this contains sufficient information for performing the task successfully.

In criminal investigations, the forming and testing of investigative hypotheses is frequently applied (Fahsing, 2016; Monckton-Smith et al., 2013) and these hypotheses should be disclosed as part of the task/mandate for the DF work (Jackson et al., 2015). This enables the practitioner to understand the remit of an investigation, which in criminal contexts can be very broad. Once investigative hypotheses are known, this allows the DF practitioner to make sense of their task in finding information, which may support or refute these hypotheses. It is important to note, that without such contexts, an investigation becomes a vague, inefficient ‘find evidence’ task, where investigative hypotheses act as a compass, guiding practitioners through the examination.

Before starting any investigation, the practitioner should also have a clear understanding of the scope of the DF work they are contracted to undertake, i.e. whether the scope is investigative advice and/or a comprehensive analysis, with interpretation of the result and evaluation of evidence, or a combination of these (Jackson et al., 2015). Investigative advice may be the aim at an early stage of the investigation, where examination of the digital evidence may provide explanations of what has happened. For example, in a case of a suspected abduction, search phrases in the suspect’s web history relating to a specific location could provide leads to where the missing person may be found. An evaluative opinion might be required at a later stage of the investigation, to consider the strength of the digital evidence. An example of such an evaluation could be*: The findings from the examination provide strong support for that the child sexual abuse images were downloaded from the internet by the suspect (hypothesis 1), and no support that the images were download by a malware program (hypothesis 2)*. To ensure transparency,

the practitioner should document all case information that was received, the hypotheses that form the basis for the examination and the scope of the task.

4. *Knowledge and experience (competence)*: Most DF work requires a combination of different knowledge components in order to carry out a comprehensive examination where both investigative, legal and technological skills are often required to ensure a high quality examination of any digital data (Sunde, 2017). Investigative skills involve decision-making psychology, inferential reasoning and dissemination/presentation skills. Investigative skills also include knowledge and experience with the particular crime phenomenon under investigation, for example who the typical offenders or victims are, or what the typical modus operandi is. Legal skills involve knowledge about relevant criminal and procedural law, and are necessary in order to understand the particular crime under investigation and the conditions that must be violated for the crime to be committed. Technological knowledge is crucial in order to predict where relevant digital traces may be found. An examination may require highly specialized technical skills, for example in relation to file systems, networked technology or encryption.

5. *Understanding the particular offence / surrounding circumstances*: Under this thread, practitioners should be reviewed for their understanding of the suspected offence, both at a legal and procedural level. In the context of legal considerations, we do not propose that the DF practitioner should be law experts, but practitioners should have sufficient knowledge to understand those acts that constitute an offence and therefore understand if information linking to these acts becomes apparent as part of their investigation.

At Checkpoint 1, the Advisor should assess, with support of the checklist in Appendix 1, whether the practitioner or team:

- Have verified the case paperwork and received exhibits/data sets.
- Have clarified the clients requirements and expectations
- Have sufficiently documented the task and the case information they have received, the hypotheses, and whether they have a clear understanding of the scope and aim of the required investigation.
- Have the necessary investigative and legal knowledge, and technological capabilities for carrying out the required investigation.

- Have made sufficient use of the available case information, whether they understand the legal concepts and evidential themes associated with the suspected offence type(s) and whether they have identified relevant legal, investigative or technological limitations or difficulties.

3.1.2 Checkpoint 2: Identification, handling, preservation and collection

Checkpoint 2 governs those processes which focus on device identification, handling, preservation and acquisition. The following avenues of discussion exist and form part of peer advice at Checkpoint 2.

1. *Device identification and triage:* Whilst not a task for all, some DF units/practitioners may be involved in the initial search for digital datasets/exhibits. It is important that this process is undertaken correctly at the first instance, as often there will be no viable second attempts, and therefore this step of peer advice is confined to those involved in these tasks. The relevant devices may be identified based on the task description, case information and investigative hypotheses of the case. The collection of evidence should only be conducted for forensic purposes if there is reason to believe there is relevant information to support or refute the hypotheses of the case. If triage decisions are being made, following some form of on-scene triage of the digital data held on the device itself (noted as 'device triage' in Horsman et al., 2014), the practitioner must be aware of the limitations of the triage tool and any triage methodology being implemented. In such situations, they cannot operate under the belief that this approach will have captured all data in all cases.

2. *Device handling procedures*: Similar to those issues noted above with device identification, acquisition procedures in most instances should be straight forward, with organisations maintaining a validated methodology which the practitioner must follow.

3. *Device acquisition*: The acquisition of data in a digital investigation and ensuring this is done correctly is of paramount importance, and, in most cases should be a straightforward and documented process. However, remote sources of data must now be considered in all cases. The storage of potentially evidential information by online services now means that the local storage of seized devices is frequently only a subset of data documenting a series of events. Access to this data may be difficult, requiring legal authority or third party compliance.

4. *Verification of acquired data set*: Verification of acquired data forms the foundation for later investigatory work. Therefore, any acquired data must be confirmed as being verified using an acceptable method and tool, and the results should be documented.

The importance of Checkpoint 2 lies with the fact that errors at this stage of an investigation can be critical. The mishandling of a device can lead to irreversible changes or damage to digital content, which may impact the perceived success of an investigation. Errors at the data acquisition phase are often irreversible, and could have major consequences for the further investigation. For example, if incomplete or erroneous datasets are acquired under the belief that they are complete, all subsequent investigatory work involving information from these datasets could be undermined. Even though Checkpoint 2 in most cases will be considered as a routine step that is straightforward to complete, it remains one of the most important to complete in both an accurate and well-documented manner.

At Checkpoint 2, the review process should also consider that it is also common practice for laboratories to operate with technician roles, who may be responsible for device handling, acquisition and verification. If this structure is in place, then a review at Checkpoint 2 performed by the Advisor may require declarations of compliance from an additional member of the organisation's staff (the technician responsible) in order to determine compliance at Checkpoint 2.

At Checkpoint 2, the Advisor should assess, with support of the checklist in Appendix 1, whether the practitioner or team:

- Have developed a sufficient plan for the search, identified and/or triaged relevant devices, and assessed the necessity of taking the devices.
- Have assessed if the devices are handled with the established procedures, and that any deviations are justified and documented.
- Have considered both locally and non-locally stored information for acquisition, provided accurate documentation for the acquisition process, justified and documented any deviations from standard procedure, and verified the acquired data set.

3.1.3 Checkpoint 3: Examination

Checkpoint 3 focuses on the examination procedures carried out and the motivation for doing so. It is important to note that Checkpoint 3 does not concern the scrutiny of any subjective interpretation and evaluation of evidential findings; this is Checkpoint 4. The processes and procedures undertaken by a practitioner are what require evaluation at this stage with the following themes in need of addressing:

a) *Understanding of case information:* The practitioner/team should understand the particular case well enough to be able to outline an examination strategy and relevant examination procedures. In essence, where the practitioner fully and correctly understands the behaviors associated with the offence under investigation, they must also develop a strategy for assessing whether such behavior is present on any of the digital exhibits being examined. For example, a hypothetical case of peer-to-peer copyright media distribution should not solely focus on the identification of any pirated media, where processes for identifying and examining suspect communication and distribution of this material may also be important (to note, this is a non-exhaustive list, simply an example for context).

b) *Examination strategy:* An examination strategy should be developed based on the case information, the investigative hypotheses, and aim of the task. This entails that the practitioner understands the particular case well, and is capable of deducing examination requirements. If the task/mandate is general, with hypotheses at an offence level, the offence hypotheses may be developed into sub-hypotheses on both activity and source levels (ENFSI, 2015a; 2015b). In comparison, activity level hypotheses are formed regarding the actions performed on the exhibit, for example ‘image ‘X’ was downloaded from the internet (or not)’, ‘file ‘X’ is deleted (or not)” or “there is communication with the victim on exhibit ‘Y’ (or not)’. ‘Or not’ is included since this constitutes the competing hypothesis. Source level hypotheses are concerned with establishing the relation between a source and a trace. Examples of a source level hypothesis are ‘The illegal activity on the computer was performed by the user account ‘UserX’ (vs by another user account)”, or Mr X used the user account UserX when the illegal activity was performed (vs someone else used the user account UserX) .

   It is important to note that the practitioner’s examination strategy will in most cases start with general ‘offence-based’ processing, influenced by what are typical traits associated with the specific offence under suspicion. Initial results will then influence ‘follow-up’

processes which over time become more specific to the offence under investigation. The hypotheses define the scope of the examination, and the practitioner should be aware that actively searching for evidence out of scope for the hypotheses under investigation might be unlawful.

c) *Examination procedures*: Based on the prediction of possible evidential traces, where they may be located and in what form, the examination procedure that is suitable to obtain the traces should be decided. The chosen procedure should be evaluated for suitability and lawfulness, to ensure that the evaluation does not violate privacy or legal authority.

The practitioner's approach to *processing* their case should be reviewed to ensure that the processing tool chosen is adequate for the task. This includes determining what the tool is capable of, what the practitioner believes it is capable of and finally, what the practitioner is trying to achieve by any chosen 'processing' process.

The practitioner's *implementation* of any chosen tool should be evaluated to ensure that the tool is suitable for the purpose. This includes examining any tool's configuration to ensure that it is set up to process data in the way that is consistent with what the practitioner wants it to do. This is important as practitioner's may mis-operate tools either accidentally, through misunderstanding of functionality or due to a lack of training.

Any *validation and testing* undertaken or relied upon by the practitioner, to ensure that the tools used in a DF process are reliable and produce valid results. Note that this is validation of the tool performance, and not a validation of results. The practitioner/team must therefore check whether the tool/version and the particular function that should be used has been tested (ENFSI, 2015a). If not, the practitioner/team should conduct testing themselves. A non-exhaustive list of questions to ask at this point include:

- Has the tool been tested?
- Have results been verified?
- What has been done that is beyond a 'push button' approach?
- Is the tool reliable?
- Has the tool been configured correctly in the first place?

- Are there any known tool errors and limitations with the tool, and are they accounted for?
- Are there mechanisms in place to report errors and prevent further usage of that tool's function?
- Did the practitioner document for which task the respective tool(s) where used?
- Are errors or limitations documented?

In most cases, *data recovery* will take place. In regards to the recovery process there are the following key themes:

- Are the correct types of data being recovered?
- Are all the potential relevant data types being recovered? (consider resident software and the types of files associated with its usage).
- Has the recovery completed with or without errors?
- What areas have been examined, and were any excluded?
- Are there any false positives or false negatives?
- Are there any limitations with carving algorithm/tool?
- Are there issues with file fragmentation or contiguous structures?

Similar to data recovery, *data parsing* will almost certainly take place. In these cases, those questions noted above also apply. Here, it is important to ensure that the process used to parse the data is reliable, and that it has been correctly used in order to obtain all available data. Any errors, limitations or reservations should be noted.

At Checkpoint 3, the Advisor should assess, with support of the checklist in Appendix 1, whether the practitioner or team

- Have sufficient understanding of the offence under investigation, and the associated behaviors that may have left digital traces.
- Have defined hypotheses at a sufficient sub-levels, and identified relevant sources for testing them.
- Have chosen adequate processing tools, implemented the tools correctly, and assessed the reliability and validity of the tool.
- Considered data recovery and data parsing, checked results for errors, assessed and documented error, limitations or reservations.

3.1.4 Checkpoint 4: Analysis, interpretation and evaluation

Both the investigatory hypotheses and aim should have formed the basis for the development of a practitioner's analysis strategy. Via Checkpoint 4, the suitability of this approach must be scrutinised. For example, the aim might be to provide *investigative advice* about who may have been involved in a particular crime. Or, the aim could be to provide an *evaluative opinion* about whether a Trojan or suspect downloaded illegal material to a device. The aim could also be to *verify other information* in the investigation, e.g. the suspect claims to have a document relevant to the investigation on the computer, and the analysis is limited to examine whether the document is there or not. This type of analysis requires minimum or no interpretation, and could therefore be subject to factual reporting (i.e. 'x' file was present on a device). It is important to take into account whether the aim is investigative advice, evaluative opinion, verification/factual result, or a combination of these, since this creates implications for any further analysis of digital data and how the outcome of this analysis should be documented.

1. *Analysis strategy:* An analysis strategy should be developed based on the case information, the investigative hypotheses and aim of the task. If the aim is to find out what has happened, strategies could be directed towards relations, timelines, functions and the evidence itself. When the analysis strategy involves forensic questions, analysis strategies would involve the use of core forensic processes, such as identification, classification, authentication, reconstruction and evaluation, and forensic activities such as interpretation and integration (Pollitt et al., 2018).

   A *relational analysis* would focus on identifying relations between entities, such as people, mail addresses, aliases, IP-addresses, telephone numbers etc. This type of analysis may help to understand the relations between the different entities (that alias X is related to e-mail address Y), hierarchical structures, and sources of digital evidence that may have been overlooked (King, 2006; Casey, 2011).

   A *temporal analysi*s will focus on the time and sequence of events. A timeline will provide oversight over what happened in which order, and may reveal information gaps. (King, 2006, Casey, 2011).

A *functional analysis* will revolve around the functions of the computer system. This analysis may shed light to whether the computer system had the necessary functionalities in place for the criminal act to be committed (King, 2006, Casey, 2011).

Evidence analysis will focus on several aspects. Probably the most important one is to establish whether there is a connection between a particular person and the evidence. The evidence analysis will explore the ownership of and access to the evidence (King, 2006). For digital evidence, this may be a complicated process - which sometimes needs to be combined with other approaches, such as suspect interviews or criminalistics. It may also shed light on the activities related to the evidence (e.g. accessing, creating, modifying) or the knowledge of the evidence (traces of hiding, protecting from others etc.) (King, 2006).

2. *Testing and validation of findings*: Not only must the practitioner test and validate their findings, they must demonstrate it. Transparency is the key concept in regards to testing and therefore a practitioner must document what they have done, how they have done it and any data generated via testing, so these processes can be reviewed. This content should be available for peer scrutiny, and to pass on and support others in the DF laboratory or unit, making further reproduction of results more efficient in the future. Organisations undertaking DF analysis should have procedures in place to support and facilitate testing of results. If not, the practitioner is faced with two choices, either to develop and document their own test methodology, or adopt existing frames, for example the Framework for Reliable Experimental Design (FRED) (Horsman, 2018b) and the Digital Evidence Reporting and Decision Support (DERDS) framework (Horsman, 2019).

3. *Interpretation and evaluation of findings*: Interpretation involves making sense of the findings in an accurate way in the light of the particular case information. Evaluation involves assessing and assigning what weight the findings have in relation to the hypotheses under consideration. Both interpretation and evaluation of findings must be demonstrated and justified in a clear and concise manner. Any analysis which involves interpretation or evaluation should adhere to the principles of balance, logic, robustness and transparency (ENFSI, 2015b).

    a. Balance involves that the findings should be evaluated given at least a pair of hypotheses.

b. Logic means to address the probability of the findings, and determine to what degree they provide support for the respective hypotheses.

c. Robustness means that the findings should be based upon sound knowledge and experience of the trace types, and that the reporting should be capable of sustaining scrutiny and cross examination (see also three pathways in DERDS, Horsman, 2019). Robustness may also be related to how thorough the hypotheses have been tested. For example, a negative finding may only be taken into account if the search for information is conducted with sufficient precision and effort.

d. Transparency involves that the process of obtaining the result is documented and presented in an understandable manner both for the Reviewer and a non-expert.

At Checkpoint 4, the Advisor should assess, with support of the checklist in Appendix 1, whether the practitioner or team:

- Have chosen an adequate analysis strategy.
- Have validated the findings and provided sufficient documentation.
- Have interpreted and/or evaluated the findings in line with principles of balance, logic, robustness and transparency.

3.1.5 Checkpoint 5: Investigative work complete?

Checkpoint 5 is a collective check of the previous four Checkpoints. It acts as a gatekeeper, preventing unfinished investigations from proceeding to the reporting and presentation stage. Essentially, Checkpoint 5 involves a critical review as to whether both the practitioner and the Advisor agree that all the necessary investigatory work is complete. If agreement is present, then the practitioner can proceed to report their findings in a written report, and submit it for peer review. Disagreements require the undertaking of any further suggested work.

For the practitioner to proceed beyond this stage, the Advisor must first '*sign-off*' the work, signifying support for the practitioner's work. Then, the Advisor must submit a 'brief', which will contain a short description of the key points in the narrative between the Advisor and practitioner across the Checkpoints, the key points of debate, any discrepancies/issues and how these were

rectified (see Appendix 2 for checklist and Advisor Brief template). The Advisor Brief will be used if it becomes necessary to conduct the dispute resolution stage (see section 3.2.2).

3.2 The Peer Reviewer Stage

3.2.1 Reporting and presentation

The reporting and presentation stage involves a peer review of the product of the investigation. This would typically include a written statement or report, but the examination notes/log may also provide valuable information for the review. Broadly, ENFSI (2015a) note that there are mainly three different types of reporting in investigations (note: the different types reporting may be combined in one report) - investigative advice, factual reporting or evaluative opinion expert evidence. The description of findings depends on the chosen approach, and the applied approach(es) should be explicated in the report. Regardless of the report being produced, the results should be presented in compliance with the principles of balance, logic, robustness and transparency (see Section 3.1.4). In addition, the ethical and legal perspectives of the case must be taken into account.

Yet before a peer review is carried out, it is necessary to consider the scope of the review according to the Peer Review Hierarchy for DF, shown in Fig. 3.

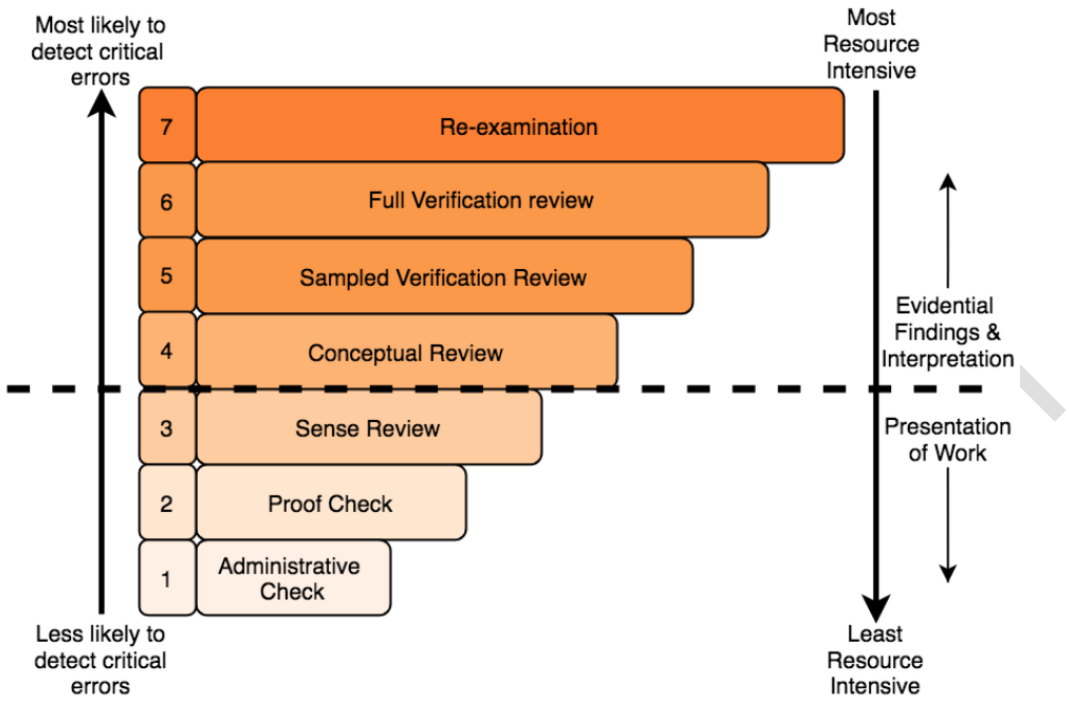


Figure 3: The 'Peer Review Hierarchy' for DF (Horsman and Sunde, 2020)

Whilst each of the levels are described below in order of rigor, it is necessary to state that each Peer Review Hierarchy level absorbs those responsibilities of the levels below it. For example, a level 4 Conceptual Review includes the specified criteria discussed below, but the Reviewer carrying out a Conceptual Review will also be expected to incorporate those requirements in the levels below (level 1-3). This should be taken into account when considering the descriptions below.

Level 1 - Administrative Check: An administrative check focuses on whether the practitioner has undertaken the correct investigation, followed clients requirements and ultimately completed those tasks on the relevant exhibits. The review of the report should focus on the following queries:

a) Is the mandate/task accurately described, and does the report cover the mandate/task and the aim of the examination sufficiently?

b) Did the practitioner/team describe the case information that was received prior to or during the examination?

c) Has the practitioner/team described their own competence in the report, and was the competence adequate for undertaking the task?

d) Is the analysis report completed in compliance with the template implemented by the unit, and is it signed?

e) Is the casefile complete?

f) Is the documentation in compliance with the standard applied in the unit?

g) Is the product in line with the agreement/contract with the client?

Level 2 - Proof Check: A proof check entails to review the document for spelling and grammar, and should focus on the following:

a) Is the report checked for spelling errors?

b) Does the grammar need to be improved?

c) If abbreviations are used, are they explained - and are they used consistently?

Level 3 - Sense Review: A sense review involves checking if the descriptions of findings and the report makes sense as a piece of evidence. The review should focus on the following:

a. Is the language understandable?

b. Are technical terms explained?

c. Has the practitioner/team clearly conveyed whether the reporting is aimed at factual reporting, provide investigative advice or an evaluative opinion (or a combination)?

d. Is it clear whether the report presents preliminary or final results?

Level 4 - Conceptual Review: A conceptual review is an extensive check of the content of the report, but without verification of the results. The main focus on the review will be on the science and logic underpinning the report. A particularly important part of this review will be to check the internal consistency between the evidence presented in the report and the conclusion. In addition to assess logical validity of the conclusion, the review should assess the report for the following elements:

1. Balance
    a. Are the hypotheses that formed the basis for the examination described?

    b. Are relevant sub-hypotheses defined at the relevant level? (activity, source)

    c. Are the results described and evaluated in relation to at least two competing hypotheses?

    d. Is there internal consistency between the results and the conclusion?

    e. Are the evaluative opinions stated in compliance with a defined structure, such as the Evidence Certainty Descriptors (Horsman, 2020) or Case Assessment and Interpretation framework (CAI) (Jackson et al., 2015), and is the chosen structure referenced?

2. Logic
    a. Are the grounds on which inferences/assumptions/interpretations are based justified and explained? Are they valid?

    b. Is the conclusion balanced, justified and explained?

    c. Does the strength of the conclusion reflect the findings it is based upon?

3. Robustness
    a. Does the examination provide sufficient basis for the conclusion?

4. Transparency
    a. Does the report refer to established processes/procedures? Are any deviances justified and documented?

    b. Are the processes and methods described accurately so that they may be repeated by others?

    c. Does the practitioner/team demonstrate a clear distinction between facts and opinions (inferences/assumptions/interpretations/evaluations about facts)?

    d. Are the findings presented accurately?

    e. Are the findings related to the context in which they were found?

    f. Are negative findings (searched, and did not find) documented (Horsman, 2018a)?

    g. Are reservations/uncertainty/limitations with methods, tools, results or conclusions conveyed?

5. Ethical/legal
    a. Did the practitioner/team conduct the examination and reporting in accordance with the criminal procedure act and applicable code of conduct/ethics?

    b. Did the practitioner/team demonstrate independence during the examination and reporting?

    c. Was the presumption of innocence operationalised during the examination and reporting?

    d. Have traces that may indicate innocence or mitigating circumstances actively been searched for and documented?

e. Have searches for information that could confirm the suspect's account been conducted and documented?

It is important to note that where a Reviewer opts for a peer review type at Levels 1-4, the review is based on the descriptions of evidential findings in the analysis report. Therefore, the Reviewer will not have the opportunity to verify any findings via these review types. These review types operate on an assumption that the practitioner has carried out all of the investigatory processes correctly up until the point of completion, where only the contents of the report is subject to evaluation as part of the review. It is not until Levels 5-7 where the review includes the verification of evidential findings, and arguably a more robust review process.

Level 5 - Sampled Verification Review: A sampled verification review includes verification of the findings uncovered by the examining practitioner. A verification of all the findings may not be feasible or necessary, and a dip-sampling approach may therefore be relevant. The Reviewer should verify a representative sample. In order to minimize interpretation errors, the Reviewer should use a different tool than the initial examiner for verification. The Reviewer should document those findings which could be verified, and which could not, and whether any deviances were uncovered. The review should focus on the following:

a. Did the practitioner/team make all findings and associated data available for verification?

b. Can a representative sample of the results be verified (by the Reviewer) through using a different tool than the initial examiner used?

c. Has the Reviewer provided accurate documentation of the verification process, and which results that have been verified?

Level 6 - Full Verification review: A full verification review involves reviewing and verifying all the findings described in the analysis report, taking into account the requirements of a sampled verification review and applying these across all data. The review should focus on the following:

a. Did the practitioner/team make all findings and associated data available for verification?

b. Can all results be verified (by the Reviewer) through using a different tool than the initial examiner used?

c. Has the Reviewer provided accurate documentation of the verification process, and which results that have been verified?

Level 7 - Re-examination: A re-examination will entail completing the examination, analysis, interpretation and reporting a second time by someone who has not worked with the case before, and has no knowledge of the results from the initial examination. The re-examination should preferably be conducted with different tools that the original examination. This will be the most time and resource intensive type of review, but will arguably provide the best opportunity for scrutiny of the practitioner's results by reducing the chance of bias. This approach may lead to new findings as well as new interpretations of existing findings.

The requirement of using another tool on level 5-7 is important, since tools may have flaws and interpret data differently. If the same tool is used for verification/re-examination, the opportunity to reveal interpretation error is lost. This may be compensated to some degree by reviewing the raw data manually, which requires expert level knowledge on file systems.

3.2.2 Dispute Resolution

The PARS model maintains an optional stage, to be called upon when required; that of *dispute resolution.* Whilst it is hoped that peer reviews take place in an amicable and supportive fashion, disagreements are almost inevitable, where in such instances, it is necessary to have appropriate procedures in place to deal with this scenario. An absence of dispute resolution runs the risk of the review process reaching a deadlock, which cannot be broken by the two parties (practitioner and Reviewer) alone. Where the Reviewer and practitioner disagree on a course of action (the need for additional work, a disagreement in evidence interpretation etc.) there must be a defined way in which this can be brought to a satisfactory endpoint.

It is important to note that the initial examiner and the Reviewer should not be left alone with the problem of solving the dispute. Any dispute resolution should be a formalised and structured

process, which should be escalated to management level. Dror and Pierce (2019, p. 7) state that “verifiers and initial examiners should not attempt to resolve the disagreement on their own, among themselves. Disagreements should be brought to the attention of management for investigation and documentation for resolution (e.g. understanding source/reason for the disagreement and soliciting additional opinions), as well as used for learning purposes.” As per documentation of dispute resolution they state: “Furthermore, the initial disagreement, as well as the outcome of the differing opinions or interpretations of the evidence, must be clearly stated in the report“ (Dror and Pierce, 2019, p 7). Those in managerial positions within an organisation should conduct and evaluation of what the dispute revolves around. They may choose to bring in a third internal Reviewer, or commission an external reviewer. It is of great importance that transparency is exercised in the dispute resolution phase. If the first examiner’s conclusions are changed, it should be clearly conveyed what the changes are, and the reasons for the adjustment. Some disputes may be complicated to solve, and it may eventually be up to the court to conduct the final assessment of upon which conclusion they base their decision.

Fig. 4 maps the proposed dispute resolution structure, which begins when the Reviewer raises an issue with a submitted report that is not agreed with by the practitioner. In this case, a dispute is raised and formal resolution procedures must begin with first engaging the management structure that a dispute is lodged. At which point, management will facilitate mediation between all parties, taking into account the Advisor’s submitted brief. Mediation is a form of dispute resolution, facilitated by a third party who supports discussions in order to determine if an agreement can be sought (Citizens Advice, 2020; British Columbia International Commercial Arbitration Centre, 2020). At this point mediation attempts to find a resolution but does not enforce or bind parties to one, as in absence of one, a third party reviewer must be appointed to undertake a blind review. Following completion of the third parties’ review, management must facilitate arbitration between all parties. Arbitration is defined as ‘a procedure in which a dispute is submitted, by agreement of the parties, to one or more arbitrators who make a binding decision on the dispute’ (WIPO, 2020). This binding decision ultimately is undertaken by those in senior managerial positions taking into account all parties involved.

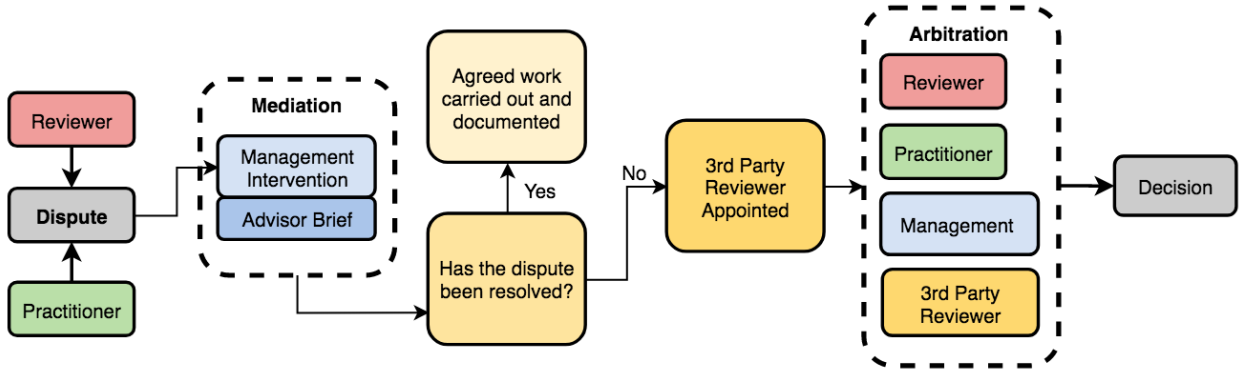


Figure 4: The dispute resolution process.

Decision: At the close of the dispute resolution phase a decision must be made. Yet what this decision is, and the impact of it, depends on the dispute itself and may differ from case to case. It is also possible that not all parties may agree with any given decisions, and where this occurs (albeit likely in the minority of cases), there must be transparency in the decision making process and the reporting of this to the client, where in some cases it may be necessary to disclose all opposing opinions. One of the main points to distinguish when making a final decision following any arbitration process, is the difference between disputes over opinion where scope for open-ended interpretation exists (for example, questions involving identifying someone on CCTV footage), and disputes over technical interpretations where it is possible to determine a right or wrong answer (an interpretation of an artefact) (UK Forensic Science Regulator, 2017). Some evidential situations may revolve around the quantity or quality of the information upon which the decision should be made, because there is necessary information to make a decision. However, sometimes the information is not sufficiently complete or accurate to reach a conclusive decision, meaning any decision could be 'inconclusive', reflecting the very limited weight of the evidence (Dror and Langenburg, 2019). For example, uncovered file fragments lacking information about where they existed on a machine prior to deletion, which user behavior they could be attributed to, or how they came to reside upon the device in the first instance may lead to an "inconclusive" result.

4 Discussion

Every DF laboratory or unit maintains their own set of challenges to consider, unique to their organisation. It would thus be naïve to believe that the PARS model could be implemented without

encountering obstacles and challenges. In this section, we will discuss issues concerning implementation and adaptability, and also strengths and limitations of the proposed PARS model. Eventually, some future directions are suggested.

4.1 How long will a review take?

One of the main assumed concerns around using PARS is the increased burden that the methodology may be considered to place upon an organisation in terms of ensuring that it is effectively resourced and implemented. It is therefore useful to consider the use of PARS as a culture change to lab environments where the review process should be costed into an investigation. This way, the overall review process becomes part of the overall investigative process, rather than becoming an element that is 'bolted-on' to the end.

The PARS model identified two key roles, the Advisor (responsible for supporting a practitioner through five Checkpoints and for the production of an Advisor Brief) and a Reviewer (responsible for the peer review of the report).

The Advisor Role: The Advisor role is continuous throughout the investigative process, and therefore we expect Advisors to operate this role alongside their own set of casework, fitting in Checkpoints in the PARS process appropriately. It is not possible to attribute an equitable amount of time to each Checkpoint, as there are differences in the size of the task at each stage, and every case maintains different challenges. Since the need for advice will vary, the Advisors should dynamically allocate their time to Checkpoints where it is most needed.

For obvious reasons, it is difficult to estimate the amount of time an Advisor will spend in total as part of this role. We put forward an estimate of 2.5 hours that the Advisor will likely spend advising across the 5 Checkpoints based on the hypothetical 'average' case. However, extraordinary situations may require more from this role. For example, if the DF practitioners normally work individually with a case, and are assigned to a case that requires teamwork, the Advisor may need to reserve additional time. There could also be technological or methodological challenges in a case that would require the practitioner to revisit the former steps of the DF process, which also would require more time from the Advisor.

The Reviewer Role: The Reviewer role is singular; to scrutinise and evaluate a practitioner's reported findings. An allocation of time here depends on the agreed review option chosen from

the Peer Review Hierarchy (Horsman and Sunde, 2020). Fig. 5 aims to provide an estimate of time allocations in a hypothetical case, and is based on the authors' own experience. The choice of review type may depend on a number of factors, and to ensure a consistent policy which a given DF laboratory or unit will commit to enforcing, the reviewer should consult the senior management when the review type is to be decided. The policy may include the adoption of different review types for volume vs serious crime types, or where a particular high profile case is under examination. In turn, a DF laboratory or unit may commit to a strategy for certain peer review types for certain offence types. Regardless of the position, it is important to consistently apply it.

7	Re-examination	Comparable to time taken in primary examination of data
6	Full Verification review	6-8 hours dependant on case size
5	Sampled Verification Review	3-5 hours dependant on case size
4	Conceptual Review	2 hours
3	Sense Review	1 hours
2	Proof Check	30 minutes
1	Administrative Check	10 minutes

Figure 5: Time allocation for review type.

Finally, it becomes difficult to put a time frame on the dispute resolution phase of PARS, but it is expected that this mechanism will only be activated in the minority of cases. It is not expected that dispute resolution will be burdensome to organisations, given its importance to quality assurance, it that it is a key and uncompromisable stage.

4.2 What if I cannot implement all of PARS?

There may be uncertainties with regards to implementation and adaptability of PARS. DF work may be automated such as in concepts like DFaaS (van Baar et al., 2014; van Beek et al., 2015; van Beek 2020), or handled as a shared task between several laboratories. The PARS model is based on the well-known and accepted DF process, with the aim of being generic enough to be implemented in any DF unit or laboratory environment, but there may be situations where the complete PARS model may not fit. The work considers all efforts to improve existing review processes as a positive step, and it is acknowledged that such wholesale adjustments to this review mechanism may be seen as disruptive if done all at once. Since the stages of PARS are described in detail, it is possible for organisations to phase PARS's implementation and to use parts of the framework, then expand to full implementation at a later stage.

The requirements for each Checkpoint and then subsequent peer review type (selected from the Peer Review Hierarchy) are fully described, and the elements that should be scrutinised in each are included within the appropriate templates found in Appendix 1 & 3. This should facilitate the implementation of PARS to be more achievable, as it allows an organization when completing the PARS paperwork to be transparent with regards to the scope of the implementation of PARS in their quality management system. A hypothetical example of a statement in the analysis report may be - '*Quality measures: following the DF process in case 'X', advice was provided at Checkpoints 1 and 2, and a Sampled Verification review of the analysis report and findings was carried out*'. In such a case, it is clear that Checkpoints 3-5 are not utilised by the organisation, and that not all findings were subject to verification during the peer review.

It may also not be possible to conduct a peer review of all the cases handled by the organisation, where in practical terms, workloads and case turnovers may exceed available peer reviewing resources. Whilst this is not ideal, there are several aspects that should be taken into account if only a selection of cases are to be subject to peer review using PARS. It may be assumed that the practitioner will adjust their operating practices, and be more careful and accurate when knowing that the work will be checked. Therefore, if only a selection of cases would undergo peer review, it is important that the case selection is done in an unpredictable manner through 'dip sampling'. To prevent cognitive bias by the Reviewer, any sample taken should include cases of different gravity, and not only include cases that contain actual findings, but also cases with no findings or inconclusive results (Ballantyne et al., 2017).

Where only a subset of cases in an organisation are to be peer reviewed, a *'risk-based approach'* for both case selection and determining the scope of a PARS implementation may be appropriate. A risk management and risk-based thinking is included in acknowledged guidelines and standards such as ENFSI (2015a) and ISO 9001:2015 (International Organization for Standardization, 2015). As part of this approach, an organisation should consider factors that will as a standard trigger a selection for peer review, as well as factors which may lead to a broader review scope (peer review types at Levels 5, 6 or 7). Hypothesised high-risk situations may include:

a) The type of case is new to the investigating practitioner, or he/she has limited knowledge of the type of case/type of data likely to reside in it.
b) The case requires the use of novel technology or techniques/methods (SWGDE, 2018).
c) The charge/indictment is mainly based on the digital evidence obtained through the DF investigation.
d) The outcome/results of the DF work may lead to severe consequences for the suspect.
e) The demographics of the case are beyond what is typically seen (for example, an abnormally large number of exhibits are included for examination).

ENFSI (2015a) underlines the relation between risk analysis/mitigation and peer-review, and suggests a combination of a dip-sampling and risk-based approach when selecting cases for peer-review.

4.3 How do you make a PARS review efficient?

There are several factors that are important for PARS to become an effective quality framework. In Fig. 5, hypothetical review times were provided, and flexibility concerning time allocation is possible. Each specific organisation must define their own variable time frameworks taking into account what works best for them.

Checklists are provided to ensure those involved in the PARS process are directed towards appropriate areas in both the Checkpoints (Appendix 1 and 2) and Peer Review Hierarchy levels of PARS (Appendix 3), and organisations adopting PARS are encouraged to use these. These checklists contain queries which the Advisor/Reviewer should seek answers to during the PARS process. Clear communication of the result of the Advisor/Reviewer's work is important, and this should be easily accessible and interpretable by those who wish to vet it. Therefore PARS uses a traditional 'traffic lights' styled-system for each Checkpoint or review level to communicate

whether the work is approved (green light), whether there are uncertainties or minor errors that should be corrected (yellow light), or finally whether there are major errors that need to be corrected (red light). The traffic light system is proposed as a simple, clear method of communicating the results of a review for distinguishing any issues which could exist at any given stage. The checklists should be available for the practitioners during casework, for self-evaluation purposes before submitting the report for peer review. The checklist will probably enable the practitioner to identify limitations and errors themselves, which they may correct before the report is handed over to the Reviewer. This will ease the burden for the Reviewer, but also for the practitioner - who will have less to correct after review is conducted.

4.4 Who can review or advise, and how?

As the PARS methodology has two distinct roles, it is necessary to consider those who are suitable candidates for each. As each role comes with their own responsibilities, it is important that an organisation thinks carefully about the assignment of these positions as they may not be best suited to everyone. In some cases, assignment of these roles will be dependent on the size of an organisation, and who is available to undertake the task. Whilst this is not ideal, it is arguably an accurate reflection on the state of the DF field, and a realistic consideration that the perfect scenario does not always exist.

In both roles, highly skilled and experienced practitioners are required, yet it should be noted that the Advisor and Reviewer roles are considered progressive. This is due to the Advisor's primary responsibility to guide and offer advice whereas the Reviewer is the gatekeeper, preventing substandard work from leaving the confines of an organisation. As such, there is a difference in the level of responsibility placed on each role. Fig. 6 provides a typical route expected to be taken by those who are involved in PARS, where those practitioners who are using PARS will start as a trainee Advisor progressing once necessary experience is gained. Therefore, those who are operational as a Reviewer have all of the necessary experience of the additional PARS roles. Organisations should also consider the development of formal training as part of the preparation for those who are to be assigned active roles in PARS. The training should aim at preparing practitioners for conducting peer review, its purposes and duties, obligations and conduct requirements expected at each stage. Both roles would require expert DF knowledge and skills, and training for the Advisor role should preferably include supervision training.

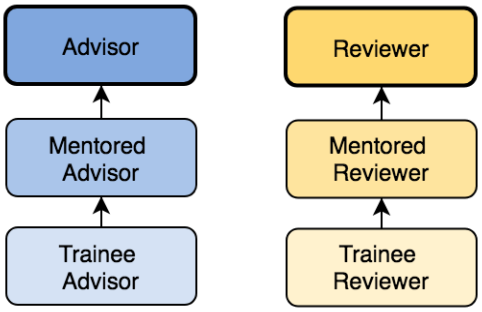


Figure 6: Role progression in PARS

The Advisor: An Advisor will provide support and guidance to the practitioner carrying out the work that is subject to review. Therefore the Advisor, at a minimum, is expected to be experienced enough to fully understand the organisational processes, have in-depth knowledge of the offence type being investigated, and investigative experience of the device types being examined and tools used to carry out this task. Therefore, where possible, it is recommended that the Advisor should be at least equal in experience to the investigating practitioner, but preferably more 'experienced'.

The Reviewer: The Reviewer role is reserved for those senior in an organisation due to the responsibility assigned to the role, and the level of knowledge and skills required from the Reviewer to perform the task. The problem this raises is that some smaller organisations may not have many practitioners employed who would be classified as suitable for this role. Where an organisation only has a small number of potential Reviewers, this may cause potential work loading issues, where all reviews may be directed consistently to a single or small number of individuals. From an organisational point of view, this may be considered acceptable as those individuals may be financially rewarded for this task, but this may inadvertently compromise the review process.

Where Reviewers are consistently performed by the same individuals, it raises concerns as to *who reviews the Reviewer*? As reviews are potentially only as robust as the knowledge of the Reviewer, it is important to ensure that reviewer work is in itself reviewed at strategic intervals to prevent bad practices and issues from going unnoticed. This acknowledges that the Reviewers themselves are not infallible, and that quality checks must occur at this stage, ensuring that the PARS process is not compromised. Those new to both the Advisor and Reviewer roles require a period of mentorship and scrutiny of their own actions. Specifically in relation to newly appointed

Reviewers, their reviews should be subject to scrutiny, analogous to a period of 'Reviewer probation'.

In all instances, the Reviewer and Advisor in any one case must be separate. This means the verification of the results may be conducted without being subject to contextual information that may cause a wide range of cognitive biases. The biases happen mainly on a subconscious level, and can not be controlled by mere willpower. First, there is a risk of 'confirmation bias', which is the tendency to look for information that is consistent with your own opinions of a given scenario and overlook or 'explain away' information that contradicts it (Nickerson, 1998). Second, escalation of commitment may occur when the Advisor has invested time and effort in the results, and may fail to see the limitations due to this (Sleesman et al., 2018). Third, a 'bias cascade effect' may also occur meaning that bias cascading from the previous Checkpoints roll into the review phase (Dror et al., 2017). Forth, a 'bias snowball effect' can occur meaning that the bias grows stronger due to several sources of irrelevant information being integrated and influencing each other (Dror et al., 2017). Fifth, there is a risk of 'status effects', which means that the perceived status of the Advisor or initial examiner biases how the result is evaluated by the Reviewer (Mattijssen et al., 2020).

As a result, it is argued that the most effective way to avoid cognitive bias implications is to ensure both the Advisor and Reviewer are separate individuals and to manage the information that the Reviewer gets access to, and which time certain information is introduced. To ensure high quality of the review, the initial examiner should not have the opportunity to appoint Reviewers themselves. "Reviewer 'shopping' may create groups with common interpretation and reporting practices within disciplines, where some groups provide conservative opinions and others more liberal interpretations of the weight of evidence. These may produce undesirable relations of trust and deference." (Ballantyne et al., 2017 p. 72).

4.5 Blind reviews

A change in Advisor/Reviewer personnel raises the question as to whether the Reviewer stage should be conducted blind. From other forensic science disciplines, there is a significant body of empirical research on the effect of contextual information on observations and decisions (see an overview in Dror, 2016). Double blind peer review is also generally accepted as the gold standard for peer review in academic journals. The effectiveness of peer review relies on that also this task is performed in the correct manner. If the Reviewer is not blinded for irrelevant contextual

information (in PARS - who performed the initial examination, who was the Advisor, what was the conclusion) there is a risk that the bias cascades from the investigation stage to the peer review stage. The result may be that the Reviewer reaches the same conclusion as the initial examiner - caused by the influence of the biasing information.

Since we cannot exclude what we already know by mere willpower, a more effective approach is to ensure that the Reviewer is unexposed to the biasing irrelevant contextual information before conducting a review. “Blind verification is true verification, not a ‘rubber stamp’, as it forces the verifier to properly examine the evidence and enables one to see if they reach the same conclusion” (Dror and Pierce, 2019 p. 7). Blind peer review in PARS entails that the Reviewer:

a. Does not know who gave advice.
b. Does not know who conducted the initial examination.
c. Does not know the initial conclusion (Peer Review Hierarchy level 4-7).

Blind peer reviewers and their examiners are more likely to disagree in the blind peer review procedures compared to when they see the others interpretation and conclusion (Mattijssen et al., 2020). Whilst disagreement does not equate to a quality review, it does help to negate ‘laboratory politics’ and/or a fear of scrutinising colleagues for fear of backlash. However, it is also acknowledged that in smaller organisations this may not be possible to avoid.

As part of a blind review at the Peer Review Hierarchy level 4-6, the Reviewer should first read the report and / or verify the results, before deciding upon the conclusion. At this point, the original practitioner generated conclusion should be unmasked, to assess whether they are consistent. If so, the result is verified. If not, the next step would be to move to the dispute resolution level (See Section 3.2.2), where the Advisor Brief document should also be taken into account. The reason for this approach is to prevent circular reasoning from conclusion to the evidence, where the conclusion affects the interpretation of results. The approach is inspired by Linear Sequential Unmasking (Dror et al., 2015), which is a procedure for managing when task relevant information is introduced in the forensic decision process in order to minimize bias.

A full re-examination at the Peer Review Hierarchy level 7 should be conducted by someone who has not worked with the case before. The re-examiner should not read the report from the initial examination, and thus have no knowledge of the results/conclusion prior to conducting the task.

In the review process, the initial examiner may also be influenced by contextual factors, such as the status or reputation of the Reviewer. If the initial examiner knows who the Reviewer was, the status effect may cause the initial examiner to accept erroneous conclusions by the Reviewer. The practitioners should thus not be informed about who conducted the peer review of their report.

Conclusively, we recommend double blind review as the gold standard also for DF peer review, although we acknowledge that this is not achievable for all units and organisations.

4.6 How to capture the PARS review

As part of this work, three template documents are provided to support those seeking to implement PARS. The checklists may be used as is, forming a paper trail of the peer advice and peer review stages of PARS, or may be digitalised and implemented in systems where parts of the DF process is automated or handled in a digital environment.

The PARS Advisors template (Appendix 1): This document is a checklist that covers all of the Checkpoints (1-4) where the Advisor will have a narrative with the investigating practitioner, and offer peer advice, including any respective queries that should be considered. In essence, every case should maintain a single PARS Advisors template, which stays with the case documentation for the duration of the investigation, and be updated at every Checkpoint by both the practitioner and Advisor with regards to any agreed actions. The PARS Advisors template stays with the case itself.

The Advisor Brief (Appendix 2): At Checkpoint 5, the Advisor should complete an Advisor Brief document, which is a short summary of the narrative had with the investigating practitioner, agreed/disagreed actions and any key points of note. The Advisor Brief document is signed by the Advisor, and is used in those cases where it is necessary to run the dispute resolution process.

The PARS Peer Review Hierarchy template (Appendix 3): This template is a checklist for the peer review stage, and covers the Levels 1-6 in the 'Peer Review Hierarchy'. When a Reviewer has determined the level of review that will take place, this should be acknowledged on the document and therefore only the corresponding review level boxes should be complete.

When a review is finished, a copy of a completed PARS Peer Review Hierarchy template should be provided to the investigating practitioner, supplying them with the results of the review, feedback and guidance on identified errors, any uncertainties which exist in the analysis report and which parts that need to be improved. The template provides oversight over what has been checked and reviewed as part of the review itself, and enables scrutiny and evaluation of the review itself. In addition, it also facilitates transparency on what has not been scrutinized, defining the scope of the review which is important to avoid any misconceptions or inflated trust with regards to the quality of any given evidence. When complete, these documents should be archived.

Archiving: As with all case data, a period of archiving is necessary in order to facilitate any future work requests or to fulfil necessary legal requirements of a specific investigating jurisdiction. The formal recording of a PARS review using the PARS template maintains three benefits to the organisation. First, a traditional peer review processes may be informal in structure and lack the rigorous recording of what took place and the outcome of a review. Whilst records of a review taking place, and by whom, may be maintained by an organisation, they may not be robust enough to be able to evaluate the depth of review carried out at that period of time. The PARS template provides a detailed structure for describing all actions and outcomes as part of the review in a format which is easily archivable, retrievable and understandable (allowing future scrutiny of the process where required). Second, a benefit of formally recording the PARS review is the ability to learn from past reviews for the purpose of continual quality improvement. (Obenson and Wright, 2013). Third, as previously stated, roles in PARS are progressive. Therefore, those who are trainee Advisors and Reviewers may find archived reviews a useful training tool, subject to having the ability to do this, taking into consideration any organizational or jurisdictional constraints.

For transparency, a practitioner's analysis report(s) should include a statement conveying whether peer review has been performed, and at which level in the Peer Review Hierarchy. In addition, any significant disagreements in opinion should be stated in the report together with a description of how a disagreement was solved, and a justification of any approach, which brought about the final result noted in the report (Ballantyne et al., 2017). Where there are any unresolved disputes regarding the content of the report, a declaration within the report should also be made.

4.7 What is the cost of implementing PARS?

Quality assurance (QA) in the form of advice, and quality control (QC) in the form of peer review, are organisational activities that will inevitably require resources. From a monetary point of view

the cost of quality will according to what is referred to as ‘Crosby’s model’ be a product of the cost of conformance with the QA and QC measures, plus the cost of non- conformance, which involves the cost of error-corrective measures (Schiffauerova and Thomson, 2006). The cost of QA and QC in PARS lie with first, training Advisors and Reviewers in order to engage effectively with the PARS process, second PARS implementation costs, ensuring the correct processes are in place to facilitate PARS becoming part of an organization's operational practices and third, time spent on PARS activities.

Any DF organisation needs to understand and acknowledge the value of QA and QC in order to invest in this activity. By implementing the proposed peer advice stage, it is expected that fewer errors and uncertainties will enter the formal peer review stage, which should lead to smaller workloads at the latter stage. The cost of implementing PARS should be considered against the potential costs which may be incurred through a lack of QA and QC measures. In such cases, inadequate measures for QA and QC may cause reputational damage, a loss of profitability (having to allocate additional resources to redo work), lead to lost market shares and in some cases have legal repercussions. From a criminal justice system perspective, undetected errors and uncertainties may lead to miscarriages of justice.

4.8 Is PARS efficient?

To our knowledge, very few digital investigation/digital forensic process models have integrated QC (see the overview offered in e.g. Casey, 2011; Yusoff et al. 2011). Of those who have integrated QC, it has been defined as a reactive activity (see for example Carrier & Spafford, 2003; Baryamereeba & Tushabe, 2004). PARS aims at improving quality prior to the peer review stage, through implementing a proactive peer advice stage, which runs alongside the natural stages of an investigation.

Traditionally, peer review has been divided in technical and administrative peer review (Ballantyne et al., 2017 Watson and Jones, 2013). The Peer Review Hierarchy for DF implemented in PARS (Horsman and Sunde, 2020) presents a more fine-grained approach to QC in the peer review stage, which enables higher transparency in the scope of the QC activities that have been performed. In addition, PARS includes the peer advice stage, which may contribute to the improved quality of any DF casework prior to entering the peer review stage.

It is important to note that peer review itself is not a foolproof concept for error mitigation, but providing it is implemented and resourced appropriately, it may contribute to improve quality (Welner et al., 2012). The error-detection rate will rely heavily on the quality of the peer review procedure, where a malfunctioning quality framework can actually have a negative effect. A poorly performed peer review will add another layer of trust to what is already often perceived as objective and reliable evidence, which may obscure the true probative value of the evidence. The aim with PARS is primarily to enhance the quality of ongoing DF investigations by preventing and uncovering errors, and second, to ensure systematic learning and improvement within DF organisations, where through systematic peer review, system errors may be identified. The systematic implementation of Checkpoints for advice would strengthen the organisation's ability to correct systematic errors. Through the support of an Advisor during the DF process, investigating practitioners have an opportunity to learn and improve their skills during the investigation. By increasing the ability to do the task correctly in the first instance, practitioners are in a position to learn from successful experiences instead of mistakes.

The quality of the result of a DF investigation should not be just a matter of trust, but one which is objectively evidencable with transparency in the QA and QC surrounding an investigation. Transparency is therefore a key factor of PARS. Instead of a guarantee of the credibility of any DF results solely being based on having possession of any appropriate accreditation or certification, transparency in those measures that have been applied (or not) during the evaluation of a given set of investigation results are seen as important. While the concepts of a technical and administrative review are used by ENFSI (2015a) and ISO/IEC (Watson and Jones, 2013) are quite general in nature, the Checkpoints, Peer Review Hierarchy and checklists in the PARS framework represent a standardised and transparent evaluative process, which makes it possible to assess in greater detail what has been done to improve the quality and control the final result of an investigation.

Those labs with ISO 17025 accreditation will have implemented a quality management system where QC procedures form one of several measures. Quality is according to ISO 9000:2015 (EN) (International Organization of Standardization, 2015) is the “degree to which a set of inherent characteristics of an object fulfils requirements”. The ‘object’ in this definition adheres to a number of manifestations such as a product, service, process, person, organisation, system or resources, which entails that quality may be measured at several levels and perspectives. Although it would be useful to compare a PARS implementation to implementation of other standards or guidelines,

we acknowledge that measuring and comparing quality and efficiency in a meaningful way is not straightforward.

There may be several advantages with implementing a quality management system such as PARS. First, correcting errors when they have cascaded further into the investigation or trial and influenced decisions can be very resource intensive. A recent example of failure to detect errors in digital evidence is the Danish Telecom-scandal, where errors in processing software were detected during the spring of 2019 (Sorensen, 2019). After investigating the causes, an external review concluded that apparent errors could be traced back to 2010. All the criminal cases with telecom data from 2010 until the time the error was corrected (more than 10,000 at the time of writing) are currently under review. The external review uncovered a virtually non-existing quality regime, and that some of the errors could have been detected with non-advanced quality checks, such as comparing the number of rows in raw data vs the processed spreadsheets (Deloitte, 2019).

By implementing PARS, the ability to detect error may increase during the Checkpoints in the peer advice stage, and the peer review stage forms an additional opportunity for error detection and correction prior to handing over the result of the DF examination. Second, implementing the peer advice stage is the opportunity to improve the knowledge and skills of the practitioners while doing their daily work. While the practitioners may gain adequate formal training through courses and education, there is no guarantee that the skills are sufficient for eliminating errors or ineffective decisions during the DF process. The Checkpoints during the peer advice stage will provide immediate feedback, which is of significant importance for learning (Vera & Crossan, 2004). Third, the peer advice stage and Advisor Brief may provide valuable information and documentation to the organisation about where the risk of errors are most present. This knowledge may be used for correcting system errors, and identify whether there is a need for additional tools, training, and/or changes in the procedures.

Implementing QA and QC measures within PARS model will come with a cost. However, it is hoped that this is seen by any organisations as an investment in ensuring the quality of their product.

4.9 Future directions

To gain knowledge about the effect of advice and review on the quality of DF investigative work, the implementation of PARS should be evaluated, and itself subject to review. In essence, this

series of work has designed and proposed PARS, and the next planned logical stage of our work is to evaluate the implementation of PARS by actual organisations. In addition, more knowledge is required regarding where in the DF process errors are most likely to occur, and which measures are best positioned to uncover or prevent these. Establishing this information will allow the refinement of the PARS review process. More information about the relationship between variables such as case type, practitioner experience and risk factors vs the scope of a review (level in the Peer Review Hierarchy) will also provide support for the continued development of effective peer review practices.

5 Conclusion

The forensic science community have universally adopted peer review, and very often in the form of verification, as an essential part of systems for quality management and error mitigation (Ballantyne et al., 2017). Whilst arguably, peer review is an important part of the DF process (as recommended in DF by organisations such as ENFSI (2015a), SWGDE (2017; 2018) and Interpol (2019)), there is no academic commentary discussing how it should be undertaken, leading to this piece being one of the first to propose and map a peer review process for this discipline.

The PARS model outlines a six stage process, the peer advice stage where five Checkpoints are undertaken via an Advisor, followed by a sixth peer review stage where the Reviewer must identify an appropriate level of peer review in line with the 'Peer Review Hierarchy'. In addition, dispute resolution is mapped and built into the process. In comparison to traditional peer review forms, which are generally considered to be a single staged approach at the close of an investigation, PARS approaches peer review with a more methodological comprehensive strategy in an effort to facilitate a robust peer review.

Whilst it is best practice for an organisation to have a peer review of all case work produced by their practitioners, it is acknowledged that this may be unachievable at this time given challenges such as growing backlogs, and constantly changing technologies. It is therefore necessary to state that PARS is not an 'all or nothing' approach to peer review. PARS is a flexible framework which can be compartmentalised and implemented over time, when an organisation finds it appropriate to engage in more of the elements it contains. It is argued that a shift to the systematic implementation of advice and review of a limited number of cases is a first step towards effective peer review.

We argue that an investment in the robust peer review methodology offered by PARS will provide long-term quality assurance benefits for an organisation. There is no doubt that reviewing work will involve a cost to the organisation in the form of time and resources, and this is acknowledged here. It requires vast knowledge not only to understand digital evidence, but also to be in a position to effectively question the evidence itself, and the trustworthiness of the process that resulted in it. We question whether there is sufficient knowledge in judicial systems to challenge the quality of digital evidence, which entails the knowledge to ask the right questions and understand the implications of the answers. Therefore, it is imperative that these checks are undertaken before any investigative work reaches this stage. Preventing errors through providing advice at the defined Checkpoints, and increasing ability to detect erroneous and misleading conclusions through implementing the Peer Review Hierarchy, is in our opinion a good investment in quality and an important measure for safeguarding the rule of law.

Acknowledgement:
We are grateful to the journal editors and anonymous reviewers whose helpful comments, suggestions and constructive criticism have improved the content and sharpened the focus of this article.

This paper was funded by a grant from the Norwegian Police University College awarded to the second author for pursuing a PhD.

References

ACPO, 2006. Murder Investigation Manual. Wyboston.

Agarwal, A., Gupta, M., Gupta, S., & Gupta, S.C., 2011. Systematic digital forensic investigation model. International Journal of Computer Science and Security (IJCSS), 5(1), 118-131.

Baryamereeba, V. & Tushabe, F. 2004. “The Enhanced Digital Investigation Process Model”, in Proceeding of Digital Forensic Research Workshop, Baltimore, MD.

Ballantyne, K. N., Edmond, G., & Found, B., 2017. Peer review in forensic science. Forensic science international, 277, 66-76. https://doi.org/10.1016/j.forsciint.2017.05.020.

Bitzer, S., Heudt, L., Barret, A., George, L., Van Dijk, K., Gason, F., & Renard, B., 2018. The introduction of forensic advisors in Belgium and their role in the criminal justice system. Science & Justice, 58(3), 177-184. https://doi.org/10.1016/j.scijus.2017.11.002.

British Columbia International Commercial Arbitration Centre, 2020. Difference between Arbitration and Mediation. Available at: http://bcicac.com/about/what-is-mediationarbitration/difference-between-arbitration-and-mediation/ (Accessed March 28, 2020).

Carrier, B., & Spafford, E. H., 2003. Getting physical with the digital investigation process. International Journal of digital evidence, 2(2), 1-20.

Casey, E., 2011. Digital evidence and computer crime: Forensic science, computers, and the internet. Amsterdam: Academic press.

Casey, E., Ribaux, O., & Roux, C., 2019. The Kodak syndrome: risks and opportunities created by decentralization of forensic capabilities. Journal of forensic sciences, 64(1), 127-136.

Casey, E., 2019. The chequered past and risky future of digital forensics. Australian Journal of Forensic Sciences, 51(6), 649-664.

Citizens Advice, 2020. Using mediation to help you separate. Available at: https://www.citizensadvice.org.uk/family/ending-a-relationship/how-to-separate/mediation-to-help-you-separate/ (Accessed March 28, 2020).

Cusack, B., 2019. Extracting Benefits from Standardization of Digital Forensic Practices. Policing: A Journal of Policy and Practice. https://doi.org/10.1093/police/paz064.

Deloitte, 2019, October 1st. Undersøkelse af Rigspolitiets håndtering av historiske teledata. Bilag 3 [Review of the Danish National Police handling of historical telecom data. Appendix 3] in Director of Public Prosecutions and Director of Danish National Police, (September 28, 2019). Redegørelse om teledatasagen [Briefing on the teledata case]. Available at: https://www.justitsministeriet.dk/sites/default/files/media/Pressemeddelelser/pdf/2019/bilag_3.pdf (Accessed March 30, 2020).

Dror, I. E., 2016. A hierarchy of expert performance. Journal of Applied Research in Memory and Cognition, 5(2), 121-127.https://doi.org/10.1016/j.jarmac.2016.03.001.

Dror, I. E., & Cole, S. A., 2010. The vision in “blind” justice: expert perception, judgment, and visual cognition in forensic pattern recognition. Psychonomic bulletin & review. Bull. Rev. 17 (2) 161–167. https://doi.org/10.3758/pbr.17.2.161.

Dror, I. E., & Langenburg, G., 2019. “Cannot decide”: the fine line between appropriate inconclusive determinations versus unjustifiably deciding not to decide. Journal of forensic sciences, 64(1), 10-15. https://doi.org/10.1111/1556-4029.13854.

Dror, I. E., Morgan, R. M., Rando, C., & Nakhaeizadeh, S., 2017. Letter to the editor—The bias snowball and the bias cascade effects: Two distinct biases that may impact forensic decision making. Journal of forensic sciences, 62(3), 832-833. https://doi.org/10.1111/1556-4029.13496.

Dror, I. E., & Pierce, M. L., 2019. ISO standards addressing issues of bias and impartiality in forensic work. Journal of Forensic Sciences. https://doi.org/10.1111/1556-4029.14265.

Dror, I. E., Thompson, W. C., Meissner, C. A., Kornfield, I., Krane, D., Saks, M., & Risinger, M., 2015. Letter to the editor-context management toolbox: a linear sequential unmasking (LSU) approach for minimizing cognitive bias in forensic decision making. Journal of Forensic Sciences, 60(4), 1111-1112. https://doi.org/10.1111/1556-4029.12805.

Du, X., Le-Khac, N.A., & Scanlon, M., 2017. Evaluation of digital forensic process models with respect to digital forensics as a service. arXiv preprint arXiv:1708.01730.

ENFSI, 2015a. Best Practice Manual for the Forensic Examination of Digital Technology, ENFSI-BPM-FOT-01. Version 01 (November 2015).

ENFSI, 2015b. ENFSI guideline for evaluative reporting in forensic science. Strengthening the evaluation of forensic results across Europe (STEOFRAE).

Fahsing, I. A., 2016. The Making of an Expert Detective: Thinking and Deciding in Criminal Investigations. (PhD Thesis). University of Gothenburg.

Horsman, G., Laing, C. & Vickers, P., 2014. A case-based reasoning method for locating evidence during digital forensic device triage. Decision Support Systems, 61, 69-78. https://doi.org/10.1016/j.dss.2014.01.007.

Horsman, G., 2019. Tool testing and reliability issues in the field of digital forensics. Digital Investigation, 28, 163-175. https://doi.org/10.1016/j.diin.2019.01.009.

Horsman, G., 2018a. “I couldn't find it your honour, it mustn't be there!”–Tool errors, tool limitations and user error in digital forensics. Science & Justice, 58(6), 433-440. https://doi.org/10.1016/j.scijus.2018.04.001.

Horsman, G., 2018b. Framework for Reliable Experimental Design (FRED): A research framework to ensure the dependable interpretation of digital data for digital forensics. Computers & Security, 73, 294-306. https://doi.org/10.1016/j.cose.2017.11.009.

Horsman, G., 2019. Formalising investigative decision making in digital forensics: Proposing the Digital Evidence Reporting and Decision Support (DERDS) framework. Digital Investigation, 28, 146-151. https://doi.org/10.1016/j.diin.2019.01.007.

Horsman, G. & Sunde, N. (2020). Part 1: The Need for Peer Review in Digital Forensics. Forensic Science International: Digital Investigation, 35, 301062. https://doi.org/10.1016/j.fsidi.2020.301062.

International Organization for Standardization, 2015. ISO 9001:2015 (EN) Quality Management Systems - Requirements.

International Organization for Standardization, 2017. ISO/IEC 17025:2017 (EN) General requirements for the competence of testing and calibration of testing and calibration laboratories.

Interpol, 2019. Global Guidelines for Digital Forensic Laboratories. Available at: https://www.interpol.int/content/download/13501/file/INTERPOL_DFL_GlobalGuidelinesDigitalForensicsLaboratory.pdf (Accessed April 1, 2020).

Jackson, G., Aitken, C., & Roberts, P., 2015. Case assessment and interpretation of expert evidence. Guidance for judges, lawyers, forensic scientists and expert witnesses. Practitioner guide No 4. Royal Statistical Society. Available at: http://www.rss.org.uk/Images/PDF/influencing-change/rss-case-assessment-interpretation-expert-evidence.pdf (Accessed April 1, 2020).

Jafari, F. & Satti, R.S., 2015. Comparative analysis of digital forensic models. Journal of Advances in Computer Networks, 3(1), 82-86. https://doi.org/10.7763/jacn.2015.v3.146.

King, G. L., 2006. Forensics plan guide. SANS Institute. 1-172.

Kohn, M.D., Eloff, M.M. & Eloff, J.H., 2013. Integrated digital forensic process model. Computers & Security, 38, 103-115. https://doi.org/10.1016/j.cose.2013.05.001.

Köhn, M., Olivier, M.S. & Eloff, J.H., 2006, July. Framework for a Digital Forensic Investigation. In ISSA. 1-7.

Mattijssen, E. J., Witteman, C. L., Berger, C. E., & Stoel, R. D., 2020. Cognitive biases in the peer review of bullet and cartridge case comparison casework: A field study. Science & Justice. https://doi.org/10.1016/j.scijus.2020.01.005.

Monckton-Smith, J., Adams, T., Hart, A.G., & Webb, J., 2013. Introducing forensic and criminal investigation. Los Angeles: Sage.

National Academy of Sciences, 2009. Strengthening Forensic Science in the United States: a Path Forward, National Academies Press, Washington, DC. https://doi.org/10.17226/12589.

National Centre for Policing Excellence, 2005. Practice Advice on Core Investigative Doctrine. Wyboston: NCPE.

Nickerson, R. S., 1998. Confirmation bias: A ubiquitous phenomenon in many guises. Review of general psychology, 2(2), 175-220. https://doi.org/10.1037/1089-2680.2.2.175.

Obenson, K., & Wright, C.M., 2013. The value of 100% retrospective peer review in a forensic pathology practice. Journal of forensic and legal medicine, 20(8), 1066-1068. https://doi.org/10.1016/j.jflm.2013.09.033.

Pollitt, M., Casey, E., Jaquet-Chiffelle, D. O., & Gladyshev, P. A., 2018. Framework for Harmonizing Forensic Science Practices and Digital/Multimedia Evidence, Organization of Scientific Area Committees for Forensic Science. https://doi.org/10.29325/osac.ts.0002.

President's Council of Advisors on Science and Technology, 2016. Report to the President: Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods, Executive Office of the President of the United States, Washington, DC.

Savage, S. P., & Milne, B., 2011. Miscarriages of justice. In: Newburn, T., Williamson, T., & Wright, A. (Eds.), Handbook of Criminal Investigation (pp. 636-653). London: Routledge. https://doi.org/10.4324/9780203118177.ch25.

Schiffauerova, A. & Thomson, V., 2006. A review of research on cost of quality models and best practices. International Journal of Quality and Reliability Management, Vol.23, No.4, 2006. https://doi.org/10.1108/02656710610672470.

Sleesman, D. J., Lennard, A. C., McNamara, G., & Conlon, D. E., 2018. Putting escalation of commitment in context: A multilevel review and analysis. Academy of Management Annals, 12(1), 178-207. https://doi.org/10.5465/annals.2016.0046.

Sorensen, M. S., 2019. Flaws in Cellphone Evidence Prompt Review of 10,000 Verdicts in Denmark. New York Times. Aug 20, 2019. https://www.nytimes.com/2019/08/20/world/europe/denmark-cellphone-data-courts.Html. (Accessed August 22, 2020).

Staw, B. M., 1981. The escalation of commitment to a course of action. *Academy of management Review*, *6*(4), 577-587.

Sunde, N., 2017. Non-technical sources of errors when handling digital evidence within a criminal investigation (Master's thesis) Norwegian University of Science and Technology. http://hdl.handle.net/11250/2450280.

SWGDE, 2017. Framework of a Quality Management System for Digital and Multimedia Evidence Forensic Science Service Providers Version: 1.0, (September 25, 2017). Available at: https://www.swgde.org/documents/Current%20Documents/SWGDE%20Framework%20of%20a%20Quality%20Management%20System%20for%20Digital%20and%20Multimedia%20Evidence%20Forensic%20Science%20Service%20Providers (Accessed April 1, 2020).

SWGDE, 2018. Establishing Confidence in Digital and Multimedia Evidence Forensic Results by Error Mitigation Analysis, Version: 2.0, (November 20, 2018). Available at: https://www.swgde.org/documents/Current%20Documents/SWGDE%20Establishing%20Confidence%20in%20Digital%20Forensic%20Results%20by%20Error%20Mitigation%20Analysis (Accessed March 31, 2020).

Tully, G., Cohen, N., Compton, D., Davies, G., Isbell, R., & Watson, T. (2020). Quality standards for digital forensics: Learning from experience in England & Wales. Forensic Science International: Digital Investigation, 200905.

UK Forensic Science Regulator, 2015. Cognitive Bias Effects Relevant to Forensic Science Examinations. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/510147/217_FSR-G-217_Cognitive_bias_appendix.pdf (Accessed March 24, 2020).

UK Forensic Science Regulator, 2017. Codes of Practice and Conduct. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/638254/128_FSR_fingerprint_appendix__Issue2.pdf (Accessed March 24, 2020).

Van Baar, R. B., Van Beek, H. M. A., & Van Eijk, E. J., 2014. Digital Forensics as a Service: A game changer. Digital Investigation, 11, 54-62. https://doi.org/10.1016/j.diin.2014.03.007.

Van Beek, H. M. A., van den Bos, J., Boztas, A., van Eijk, E. J., Schramp, R., & Ugen, M., 2020. Digital forensics as a service: Stepping up the game. Forensic Science International: Digital Investigation, 35, 301021.

Van Beek, H. M. A., van Eijk, E. J., van Baar, R. B., Ugen, M., Bodde, J. N. C., & Siemelink, A. J., 2015. Digital forensics as a service: Game on. Digital Investigation, 15, 20-38. https://doi.org/10.1016/j.diin.2015.07.004.

Vera, D., & Crossan, M., 2004. Strategic leadership and organizational learning. Academy of management review, 29(2), 222-240. https://doi.org/10.5465/amr.2004.12736080.

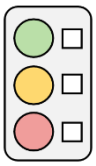
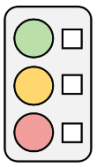
Watson, D. L., & Jones, A. (2013). Digital forensics processing and procedures: Meeting the requirements of ISO 17020, ISO 17025, ISO 27001 and best practice requirements. Amsterdam: Elsevier.

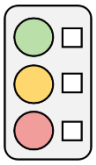
Welner, M., Mastellon, T., Stewart, J.J., Weinert, B. & Stratton, J.M., 2012. Peer-reviewed forensic consultation: Safeguarding expert testimony and protecting the uninformed court. Journal of forensic psychology practice, 12(1), 1-34. https://doi.org/10.1080/15228932.2011.588526.

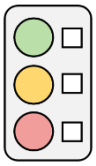
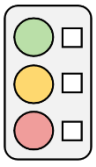
WIPO, 2020. What is Arbitration?. Available at: https://www.wipo.int/amc/en/arbitration/what-is-arb.html (Accessed March 28, 2020).

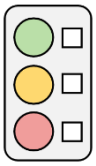
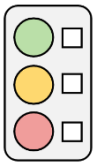
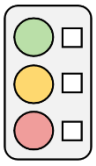
Yusoff, Y., Ismail, R., & Hassan, Z., 2011. Common phases of computer forensics investigation models. International Journal of Computer Science & Information Technology, 3(3), 17-31. https://doi.org/10.5121/ijcsit.2011.3302.

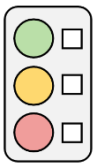
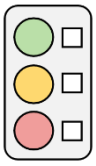
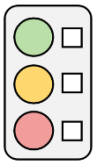
Appendix 1 - PARS Advisors template:

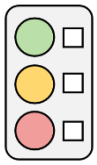
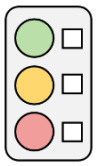
Checkpoint 1	Assessment and preparation	Advisor comment
	1. *The task/mandate and case information:* a. Has the practitioner/team provided detailed documentation regarding the case information received, together with the task to be undertaken? b. Has the practitioner/team documented the hypotheses which form the basis of the further examination? This includes for example, whether an innocence hypothesis is defined? c. Does the practitioner/team have a clear understanding of the scope and aim of the task? This may be a simple yes or no declaration ranging to a written statement by the practitioner, which describes their interpretation of what they believe they need to do.	 Comment:
	2. Knowledge and experience (competence): a. Does the practitioner/team have the necessary *investigative knowledge*, so that they are capable of predicting where relevant traces to a given case type/crime phenomenon may be found? b. Does the practitioner/team have the necessary *legal knowledge*, to understand what legal conditions must	 Comment:

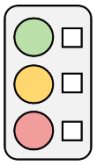
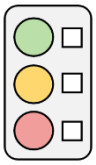
	be fulfilled to prove/disprove the offence(s) under investigation? c. Does the practitioner/team understand which acts that would be recognised as evidential, and the digital data types which may describe if such activity has occurred on a given device? d. Does the practitioner/team possess the *technological capabilities* for carrying out the required investigation? This issue is likely to occur in more specialist cases, but it is necessary to identify early in the process as to whether limitations in capability may prevent a thorough examination of available data. This could be due to limitations or knowledge, or a lack of specific devices, equipment or software.	
	3. Understanding the particular offence / surrounding circumstances: a. To what degree has the practitioner/team made use of the available case information and predicted case specific traces which are considered to be capable of shedding light on a particular suspected offence under investigation? b. Does the practitioner/team understand the legal concepts associated with the suspected offence type(s), and the relevant evidential themes that should be covered in the examination? This entails understanding the basic legislative applicability and associated actions and digital data types which may	 Comment:

	support or refute applications of the particular case under investigation. c. Has the practitioner/team identified relevant limitations or difficulties (legal, investigative, technological) of an investigation which may be apparent given details of the surrounding circumstances of the case and available digital data? Examples may include data stored in another jurisdiction (legal challenge), the offence being a novel crime phenomenon (investigative challenge) or non-locally stored data requiring further work or enquiries to secure access to it (technological challenge). It is important in such cases that there is a recognition of the existence of such data, and that any investigation does not simply proceed on the basis that it does not exist.	
	4. Client requirements and expectations: a. Has the practitioner/team clarified the scope of the work with the client, and is it achievable?	 Comment:
	5. Case paperwork and data set verification: a. Has the practitioner/team determined the correct point of contact for the task, agreed on a contract for work, with an achievable remit? b. Has the practitioner/team controlled any received exhibits/data sets?	 Comment:

Checkpoint 2	Identification, handling, preservation and collection	
	1. *Device identification and triage:* a. Has the practitioner/team made a sufficient plan for the search? b. Has the relevant devices been identified and/or triaged, and is it plausible that the device will shed light on the case under investigation? c. Is there a reasonable belief in the necessity to take/omit taking devices?	 Comment:
	1. *Device handling procedures* a. Are the devices handled in compliance with established procedures? b. If the handling of devices deviates from standard procedure, are the actions justified and accurately documented?	 Comment:
	2. *Device acquisition*: a. Has the practitioner/team provided accurate documentation describing the acquisition process? b. If the acquisition deviates from standard procedure, are the actions justified and accurately documented?	 Comment:

	c. Did the practitioner/team consider non-locally stored information, and take the necessary steps to check the availability, and if possible secure and acquire this information?	
	3. *Verification of acquired data set*: a. Has the practitioner/team verified the acquired data set?	 Comment:
Checkpoint 3	**Examination**	
	1. Interpretation of case information: a. *Does the practitioner/team fully understand the offence under investigation, and the associated behaviors that may have left digital traces?*	 Comment:
	2. Examination strategy: a. Are the investigative hypotheses of the case broken down to a sufficient sub-level in order to guide the examination of the evidence? b. Has the practitioner/team identified where the sources of information may be in order to test the sub-level hypotheses?	 Comment:

	3. *Examination procedures:* a. Has the practitioner/team chosen an adequate approach and tool(s) for processing? b. Has the practitioner/team implemented the tool(s) correctly considering the aim of the processing? c. Has the practitioner/team undertaken relevant measures to ensure that the tool is reliable and produces valid results? d. Has the practitioner/team considered to perform data recovery? If so, have the results been checked for errors? Are errors, limitations or reservations noted?	 Comment:
Checkpoint 4	**Analysis, interpretation and evaluation**	
	1. *Analysis strategy:* a. Has the practitioner/team chosen an adequate analysis strategy?	
	2. *Testing and validation of findings*: a. Have the findings been validated, and is this sufficiently demonstrated in the documentation?	 Comment:

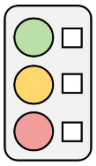
	3. *Interpretation of findings:* *a. Are the interpretations of the findings done in line with the principles of balance, logic, robustness and transparency?*	 Comment:
Checkpoint 5	**Investigative work complete?**	
	a) Is all the investigative work complete?	 Comment:

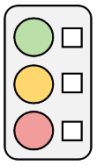
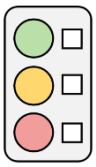
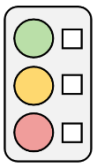
Appendix 2 - PARS Advisor Brief template

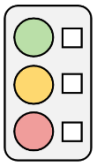
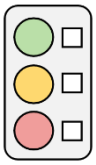
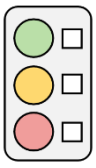
Case ID:_____ Peer Advisor:_____
Key points in the narrative between Advisor and practitioner across all four Checkpoints (e.g. if the advice led to significant change in the planned approach):
Key points of debate:
Discrepancies/issues, and how these were rectified:
Date:_____ Signature, Advisor: _____

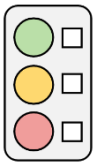
Appendix 3

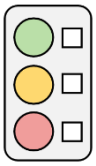
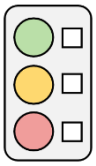
PARS Peer Review Hierarchy template

Level in the Peer Review Hierarchy	***Queries***	***Reviewer Comment***
Administrative Check	a) Is the mandate/task accurately described, and does the report cover the mandate/task and the aim of the examination sufficiently? b. Did the practitioner/team describe the case information that was received prior to or during the examination? c. Has the practitioner/team described their own competence in the report, and was the competence adequate for undertaking the task? d. Is the analysis report completed in compliance with the template implemented by the unit, and is it signed? e. Is the casefile complete? f. Is the documentation in compliance with the standard applied in the unit? g. Is the product in line with the agreement/contract with the client?	 Comment:

Proof Check	a. Is the report checked for spelling errors? b. Does the grammar need to be improved? c. If abbreviations are used, are they explained - and are they used consistently?	 Comment:
Sense Review	a. Is the language understandable? b. Are technical terms explained? c. Has the practitioner/team clearly conveyed whether the reporting is aimed at factual reporting, provide investigative advice or an evaluative opinion (or a combination)? d. Is it clear whether the report presents preliminary or final results?	 Comment:
Conceptual Review	*Balance:* *a. Are the hypotheses that formed the basis for the examination described?* b. *Are relevant sub-hypotheses defined at the relevant level? (activity, source)* c. *Are the results described and evaluated in relation to at least two competing hypotheses?* d. *Is there internal consistency between the results and the conclusion?*	 Comment:

	e. *Are the evaluative opinions stated in compliance with a defined structure, such as the Digital Evidence Certainty Descriptors (Horsman, 2020) or Case Assessment and Interpretation framework (CAI) (Jackson et al., 2015) and is the chosen structure referenced?*	
	Logic: a. *Are the grounds on which inferences/assumptions/interpretations are based justified and explained? Are they valid?* b. *Is the conclusion balanced, justified and explained?* c. *Does the strength of the conclusion reflect the findings it is based upon?*	 Comment:
	Robustness: a. *Does the examination provide sufficient basis for the conclusion?*	 Comment:
	Transparency: a. *Does the report refer to established processes/procedures? Are any deviances justified and documented?* b. *Are the processes and methods described accurately so that they may be repeated by others?*	 Comment:

	c. *Does the practitioner/team demonstrate a clear distinction between facts and opinions (inferences/assumptions/interpretations/ evaluations about facts)?* d. *Are the findings presented accurately?* e. *Are the findings related to the context in which they were found?* f. *Are negative findings (searched, and did not find) documented (Horsman, 2018a)?* g. Are reservations/uncertainty/limitations with methods, tools, results or conclusions conveyed?	
	Ethical/legal: a. *Did the practitioner/team conduct the examination and reporting in accordance with the criminal procedure act and applicable code of conduct/ethics?* b. *Did the practitioner/team demonstrate independence during the examination and reporting?* c. *Was the presumption of innocence operationalised during the examination and reporting?*	 Comment:

	d. *Have traces that may indicate innocence or mitigating circumstances actively been searched for and documented?* e. *Have searches for information that could confirm the suspect's account been conducted and documented?*	
Sampled Verification Review	a. Did the practitioner/team make all findings and associated data available for verification? b. Can a representative sample of the results be verified (by the Reviewer) through using a different tool than the initial examiner used? c. Has the Reviewer provided accurate documentation of the verification process, and which results that have been verified?	 Comment:
Full Verification Review	a. Did the practitioner/team make all findings and associated data available for verification? b. Can all results be verified (by the Reviewer) through using a different tool than the initial examiner used? c. Has the Reviewer provided accurate documentation of the verification process, and which results that have been verified?	 Comment: